BRAULIO EUGENIO CAMACHO LOYA

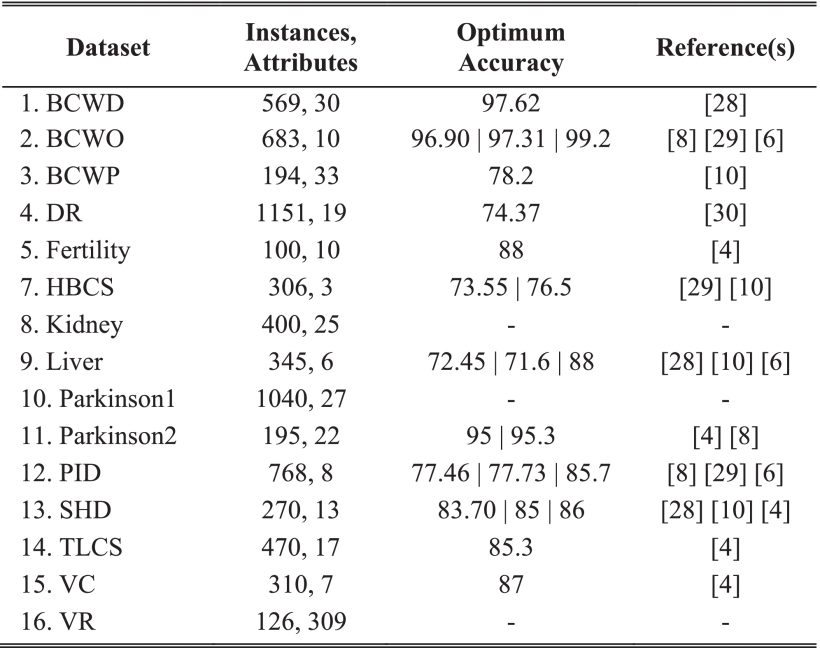
[nombre de la empresa]

Evidencia 1

Braulio Eugenio Camacho Loya 203412



Tal y como se detalló en la metodología, el primer paso para llevar a cabo el experimento es obtener los datasets públicos, relacionados con temas médicos, que recopilan la información de múltiples pacientes, con diferentes padecimientos, usados para clasificación supervisada.



Dichos datasets, son los mismos que se usan en el ensayo a comparar

A continuación, se listan los datasets usados en el experimento, acompañados de su respectiva liga de descarga.

1. Breast cancer Wisconsin diagnostic (BCWD); https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic)

2. Breast cancer Wisconsin original (BCWO);

https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(original)

3. Breast cancer Wisconsin prognostic (BCWP);

https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Prognostic)

4. Diabetic retinopathy from U. of Debrecen (DR);

https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen+Data+Set

5. Fertility (Fertility);

https://archive.ics.uci.edu/ml/datasets/Fertility

6. Haberman’s survival after surgery for breast cancer (HBCS);

https://archive.ics.uci.edu/ml/datasets/haberman's+survival

7. Chronic kidney disease (Kidney);

https://archive.ics.uci.edu/ml/datasets/chronic\_kidney\_disease

8. Liver disorders (Liver);

https://archive.ics.uci.edu/ml/datasets/liver+disorders

9. Parkinson speech (Parkinson1);

https://archive.ics.uci.edu/ml/datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings

10. Oxford Parkinson’s disease detection (Parkinson2);

https://archive.ics.uci.edu/ml/datasets/parkinsons

11. Pima Indians diabetes (PID);

N/D

12. Statlog heart disease (SHD);

http://archive.ics.uci.edu/ml/datasets/statlog+(heart)

13. Thoracic surgery survival after surgery for lung cancer (TLCS);

https://archive.ics.uci.edu/ml/datasets/Thoracic+Surgery+Data

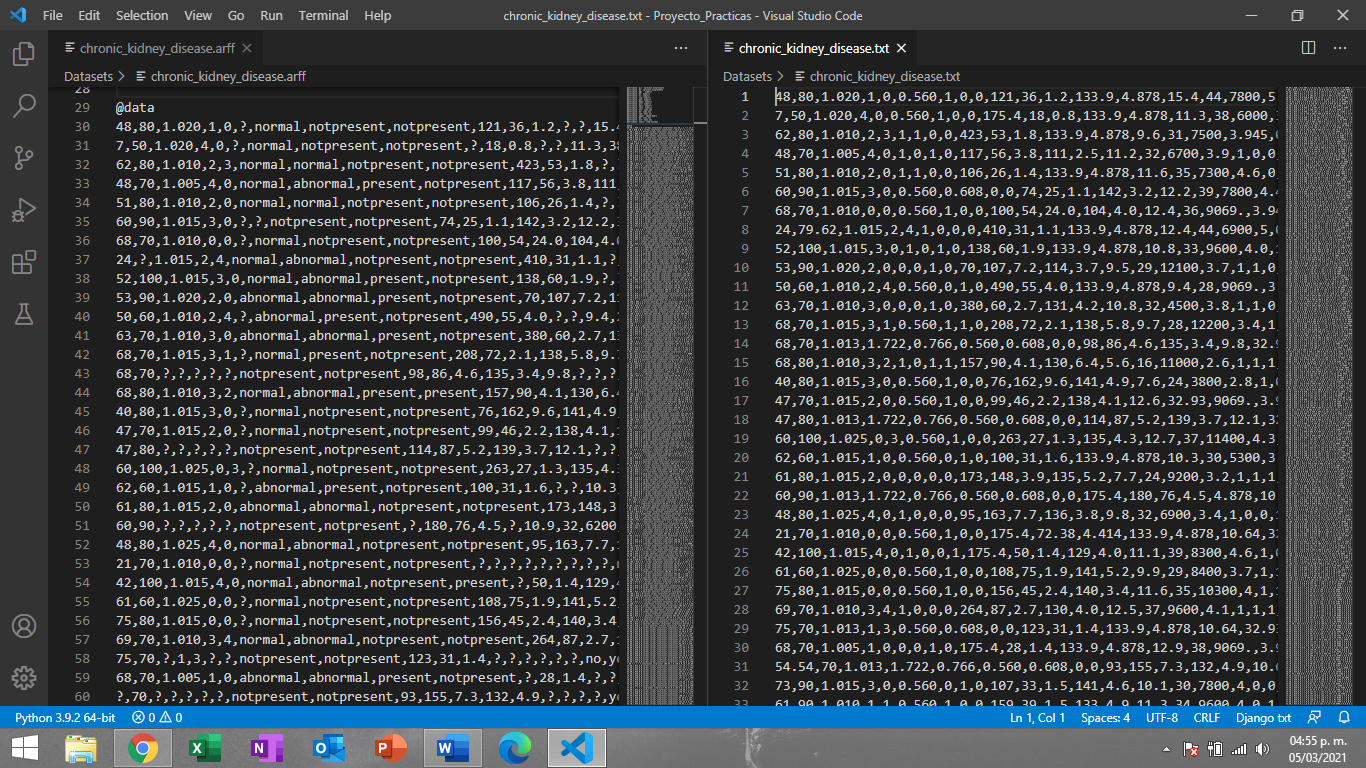
14. Vertebral column (VC);

http://archive.ics.uci.edu/ml/datasets/vertebral+column

15. LSVT voice rehabilitation (VR).

<https://archive.ics.uci.edu/ml/datasets/LSVT+Voice+Rehabilitation>

A continuación, se les realizó un proceso de limpieza, para eliminar espacios vacíos, y darles el formato necesario para el algoritmo a usar, cambiando los valores cualitativos, por cuantitativos, para facilitar su tratamiento.

El primer ejemplo, es una captura de pantalla del dataset llamado: “Chronic kidney disease” que contaba con múltiples atributos cualitativos, y espacios vacíos, los cuales fueron rellenados calculando la media de ese atributo por clase, y asignándolo.

|  |  |
| --- | --- |
| Original | Modificado |

En el segundo llamado “Thoracic surgery survival after surgery for lung cancer”, se llevó a cabo un tratamiento parecido al anterior ejemplo, cambiando los valores cualitativos, por cuantitativos

|  |  |
| --- | --- |
| modificado | Original |

